On-board Data Mining

Steve Tanner, Cara Stein, and Sara J. Graves

Networks of remote sensors are becoming more common as technology improves and costs decline. In the past, a remote sensor was usually a device that collected data to be retrieved at a later time by some other mechanism. This collected data were usually processed well after the fact at a computer greatly removed from the in situ sensing location. This has begun to change as sensor technology, on-board processing, and network communication capabilities have increased and their prices have dropped.

There has been an explosion in the number of sensors and sensing devices, not just around the world, but literally throughout the solar system. These sensors are not only becoming vastly more sophisticated, accurate, and detailed in the data they gather but they are also becoming cheaper, lighter, and smaller. At the same time, engineers have developed improved methods to embed computing systems, memory, storage, and communication capabilities into the platforms that host these sensors. Now, it is not unusual to see large networks of sensors working in cooperation with one another. Nor does it seem strange to see the autonomous operation of sensor-based systems, from space-based satellites to smart vacuum cleaners that keep our homes clean and robotic toys that help to entertain and educate our children.

But access to sensor data and computing power is only part of the story. For all the power of these systems, there are still substantial limits to what they can accomplish. These include the well-known limits to current Artificial Intelligence capabilities and our limited ability to program the abstract concepts, goals, and improvisation needed for fully autonomous systems. But it also includes much more basic engineering problems such as lack of adequate power, communications bandwidth, and memory, as well as problems with the geolocation and real-time georeferencing required to integrate data from multiple sensors to be used together.

Given the limitations of current systems, what is driving the push to develop sensor networks and autonomous systems? What place does data mining have in such environments? What techniques and solutions are people using to work around

S. Tanner (✉)
University of Alabama in Huntsville, AL, USA
e-mail: stanner@itsc.uah.edu

M.M. Gaber (ed.), *Scientific Data Mining and Knowledge Discovery: Principles and Foundations*, DOI 10.1007/978-3-642-02788-8_13,

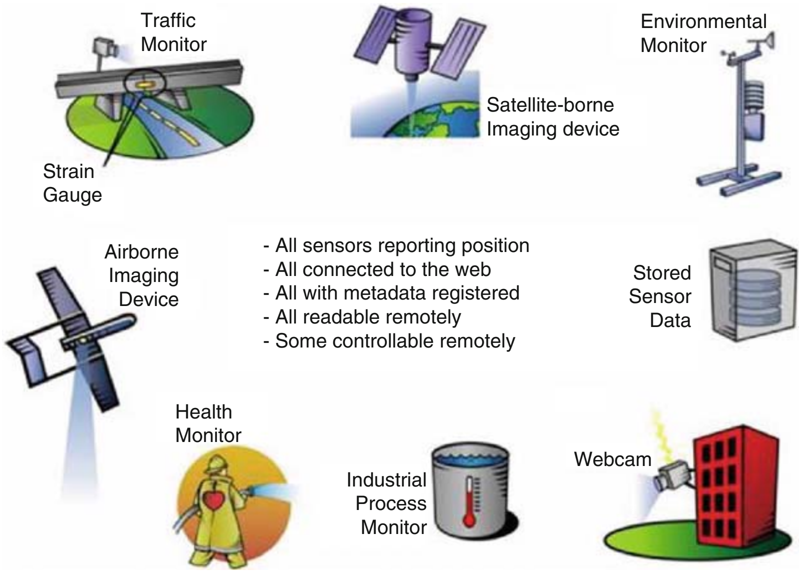

Fig. 1 The Open Geospatial Consortium's Sensor Web Enablement (SWE) specifications enable developers to make all types of sensors, transducers, and sensor data repositories discoverable, accessible, and useable via the Web (Courtesy of Open Geospatial Consortium, Inc.)

the problems encountered in the real world? At this point in time, the answers are almost all pragmatic in nature rather than absolute or rigorously and mathematically proved. Like most robotic and remote sensing systems, on-board data mining is more of an engineering endeavor than an academic one.

For on-board data processing efforts, it is important to work closely with domain scientists to verify that the algorithms accurately identify the phenomena of interest to them. They can also help with identifying false alarms or missed detections, and weighing the costs of each. Also, on-board algorithms don't have to duplicate all of the details that would result from the complete processing of an image on the ground. For example, in an Earth-monitoring application, rather than giving detailed results about a dust storm, the on-board algorithm can simply detect the dust storm and mark the data as high priority for downlink. Then, further processing can be done on the ground. Using the simplest processing techniques possible is a good approach, given the constraints. Complexity should be added only as necessary and feasible [11].

1 Problems Encountered with On-Board Mining

Before delving into current and planned efforts in on-board data mining, it is useful to consider the types of problems and issues that are likely to arise with such systems. What technical roadblocks have stood in the way of widespread use of data analysis and mining in on-board systems? The simple answer is severe resource constraints. On-board processing by definition means working in a constrained environment of some sort. The constraints vary considerably but generally fall into several broad categories: power, bandwidth, computation, and storage. In addition, the on-board environment introduces two problems associated with the data itself that the system must overcome: noise, and incorrect and incomplete measurements. What's more, the data must be georeferenced in some way, so that analysis algorithms can correlate data coming either from different sensors or from temporally separate readings. Typically in non-on-board systems, the data cleaning, and georeferencing steps are completed at a central location as part of post-processing, and result in several different data products, which are often not available until days, weeks, or even months after the data is initially gathered. This luxury is not available for on-board applications, which must deal with the noisy data results and with tagging the data through the use of a georeferencing scheme to make the data immediately usable.

1.1 Power

Perhaps the most obvious problem in on-board data mining applications is power. Many on-board mining applications are associated with sensor platforms that are physically removed from any consistent or sizable power source. This has a profound impact on how much data analysis can take place on-board the platform. Even if there is sufficient computational capability and storage, there may not be enough energy to drive the computers. This is a very different environment from desktop, cluster, or grid computation, where power is rarely a concern, and when it is, it is addressed well beforehand.

In the case of space-based systems, the lack of adequate power is due to being located in a remote and extremely hostile environment. Most space-based systems have batteries and some sort of recharging capability (e.g. fuel cell or solar), but it is rarely enough to drive all of the on-board systems all of the time. In the case of solar power generation on satellites, great care must be taken to schedule tasks such that they are compatible with the day–night orbital path of the platform – something which few data mining researchers have to contend with.

For many ground-based systems, the problem is often one of size/weight constraints. Often, the sensors must be small and disconnected from a power grid, making the use of battery power a necessity. This may be due to the remote locations of placement (e.g. for environmental monitoring in remote geographic areas). It may also be due to the need for stealth, as in the case of surveillance and military sensors, or at least discretion, as is often the case for security cameras.

The implication of this dearth of energy is the need to husband power consumption very aggressively. For data mining, that may impact the type of algorithms used, how they are used, or when they are used. Often software developers will need to consider the efficiency of a process in light of the sensor data size to determine how much energy a given algorithm will need. In some cases, the data mining plans may need to either use a lower resolution data feed (for example, sampling every other pixel instead of every pixel) or a lower sampling rate (sampling every other image instead of every image). In the most adaptive systems, the scheduler may be able to decide how to approach power consumption in a dynamic way – for example using higher power consuming algorithms during the daylight hours and a simpler, lower power version at night.

1.2 Bandwidth

Bandwidth is a significant obstacle in on-board data mining. Communication from a sensor platform, either with other sensors or with central downlink locations, is an expensive endeavor. It is expensive in terms of time, required hardware (and thus weight), power, and computational capabilities. Most developers of sensors and sensor platforms strive to limit either the frequency of communications or the amount of data sent, or both. For many sensor systems, there simply isn't enough bandwidth to send all of the gathered information, or just as limiting, there isn't enough power to drive the data through the network connections. This is especially true of wireless sensor networks, where passing information to a central processing center requires multiple hops from sensor platform to sensor platform.

This limits the amount of data mining that can be accomplished, since data mining algorithms may not have access to all of the potential data available. Data mining in pursuit of data fusion using multiple sources becomes problematic. However, data mining may actually be able to play a vital role in resolving this bandwidth problem. If the proper algorithms are employed, quick data analysis on-board a platform may be able to filter out unnecessary data, leaving only valuable data to vie for the limited communication resources.

In the case of earth science satellites, often simple filters are employed to eliminate poor sensor measurements. Cloud masks are an example of this, for instruments that can neither measure through clouds nor gather useful data about the clouds themselves. However, developing real-time cloud mask filters is nontrivial. Sun glint, snow packs, and other environmental factors make this a difficult task.

Data mining classification algorithms are a prime example of techniques that may be able to reduce the amount of data needed to be transmitted. These can even be used to deal with concept drift and environmental changes, for example through the use of a dynamically changing ensemble of classifiers, each tuned to a different environmental condition.

One must consider the impact of data mining on the need for bandwidth, and care must be taken not to overwhelm the network capacity. The most appropriate mining algorithms then are ones that reduce either the amount of data or the dimensionality. Luckily, these are often the very algorithms used in the initial data processing for data mining.

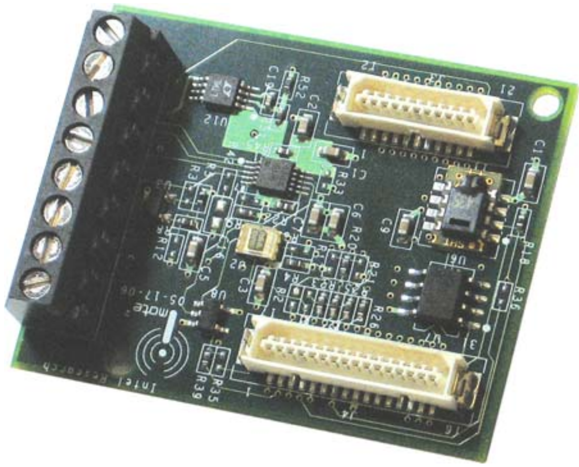

Fig. 2 Many commercial venders are offering a variety of small, inexpensive wireless sensor platforms and sensor packages. This small platform from Crossbow Technology, Inc. contains an accelerometer, temperature, humidity, and light sensors – perhaps useful for environmental monitoring applications (Image courtesy of Crossbow Technology, Inc.)

1.3 Computation

Computational limitations, combined with the need for timely results, are probably the biggest hurdle to overcome in the deployment of on-board data mining applications. Simply put, there may not be enough computing cycles to thoroughly analyze the gathered data on-board the sensor platform. This means, sacrifices must often be made in the amount or type of analysis that can fit within the computational constraints of the on-board system. Two typical approaches are to limit the resolution of the data to be analyzed or to limit the frequency of the analysis.

Consider a frame-based camera system linked with an on-board image process analysis application, in which the camera generates a series of M images per second, each of size $n \times n$ pixels. For the sake of simplicity, consider the use of a straightforward convolution algorithm as one step in the data mining process. A typical convolve algorithm uses a mask of some sort to filter the 2D image in one pass – so its computational requirement is approximately $(n^2 \times M) \times C$, where C is a constant based on the complexity of the algorithm. If there isn't enough computational power to meet this requirement, the developer has two basic choices: either limit the resolution of the images (reduce the size of n) or limit the number of images to process (reduce the size of M). The choice will depend on the application's overall goals and requirements. Which sacrifice has the least impact on the final outcome – lower resolution or fewer frames? In an application that monitors for intrusion detection, perhaps the frame rate isn't particularly important, but higher resolution images, which may help with identification of the intruders, are. In an application that is part of a rocket propulsion health management system, perhaps the frame rate and quick response to millisecond changes is the overriding concern – in this case, changes are of a fast but relatively coarse nature.

1.4 Storage and Memory

The idea of on-board data mining implies access to the data being collected. But how much data is gathered depends heavily on the sensor platform and the storage and memory resources available. Often, when the platforms are designed, these areas are given short shrift. Memory utilizes precious power, and storage takes up valuable space and weight. This may greatly impact the types of algorithms that can be used and the approaches taken to perform the data analysis. Typically, this means that there is little or no data archived, so there is little chance for multiple passes over the data. This makes taking advantage of temporal analysis difficult. Algorithms that focus on dealing with streaming data may be the best approach – many of these are designed as one-pass analysis tools.

Another issue that may arise is concept drift or changes due to environmental conditions. Since there may be no archived data to provide historical context, care must be taken to deal with such things as diurnal, seasonal, or weather-based changes. Imagine an outdoor sensor monitoring a city park. The environment there will look quite different depending on the season – snow, deciduous trees, even differing clothing fashions will shift from season to season. If an historical archive of these changes is available, adapting to them is considerably easier. But what of an autonomous system without access to such archives? For systems that have periodic contact with a base station, this issue may be addressable through manual means – by uploading new classification algorithms as seasons change, for example. Without that access, the system needs to be able to adapt to these changes on its own. Such adaptability can be especially difficult to implement. For example, one approach is to have a group of algorithms to select from, and base the selection on the latest results – if the last run of the analysis indicated rain, then use algorithms that take that into account, until rain is no longer detected. However, in a rapidly changing environment this approach may lead to a system that is constantly chasing what happened in the immediate past rather than what is happening right now.

1.5 Georeferencing

One can image a fully functioning sensor network comprised of a group of heterogeneous sensor platforms connected together via some communications means. Each of these platforms would have one or more sensors connected to it and have a computing facility including a processor and memory. In many cases, the topology of the network, and the capabilities of each platform and its sensors would be known, allowing for the development of applications geared specifically for such a network.

However, in many cases, such a priori information may not be on hand. How then can the best use of these resources be made?

One approach is to use some type of formal descriptions as a means to pass information between sensors, the platforms they are connected to, and remote computing locations separate from the sensors and platforms themselves. Through the use of such formal means, sensing processes running on the sensors or sensor platforms could advertise their capabilities while other analysis and decision support processes running at any number of locations could discover these capabilities and match them with their own needs.

2 The Use of Standards for On-Board Data Mining

There are a number of approaches being researched and implemented for resource discovery and resource advertising within computer network systems. Most of these are aimed at either specific types of services or specific types of network architectures. However, with the advent of web services protocols, there is a move to make resource discovery and advertisement a more general purpose service within the wider network community. One such approach that has direct bearing on on-board data mining is undertaken by the developers of *SensorML.*

2.1 SensorML

SensorML is part of a larger effort by the Open GIS Consortium (OGC) to develop technologies and standards that will support better use of Geographic Information Systems (GIS) data (OGC). SensorML operates under the assumption that there are three fundamental types of information that may be observed. These types of information are the object to be observed, which must have physical properties that can be measured and quantified; data values taken from observations of that object by a sensor; and meta-data about the sensor, including the location and time of the observations. Meta-data about the sensor may also include characteristics of the sensor that can help a user understand the values of the measurements as well as their quality and veracity.

SensorML is concerned primarily with the description of sensor characteristics. The information is seen by the authors [SensorML] as being used in three primary ways:

- To process raw sensor data into usable measurements and for georegistration
- To provide a limited amount of data conversion on-board the sensor or sensor platform itself
- To provide either on-board or downstream processes with sensor characteristics that may impact further data analysis.

SensorML includes descriptions of:

- Observation characteristics such as physical properties measured, quality characteristics, and response characteristics.
- Geometry characteristics such as size, shape, and spatial weight function of samples, as well as geometric and temporal characteristics of sensor data collections
- Description and documentation such as history and reference information

In summary, the primary use for the language is to make access to sensor data more automated and straightforward by supporting the description of sensors and systems of sensors.

2.2 Describing Sensors in SensorML

The most basic definition assumed by SensorML is that "sensors are devices for the measurement of physical quantities." This is clearly a very broad definition, encompassing everything from the simplest thermometer to the most complex on-orbit satellites. In fact, this also includes the use of humans as the measurement device. A more subtle aspect is that the term "measurement" is also broadly used here, and is not limited to a numeric quantity.

From a modeling point of view, sensors are typically thought of as one of two primary types. The first is in-situ sensors, which typically measure something within its immediate area (e.g. a room-monitoring system). The second is remote sensors, which typically measure things from a much greater distance (e.g. a satellite that measures reflected radiation from the Earth's surface).

This distinction between the types is important to understand because SensorML uses only one type of coordinate system to describe both types of sensors. Any geometric properties described within the schema are defined in terms of the local coordinate system (local to that sensing component). It is only through the use of an association between the sensor and its platform (with its own coordinate system), and that platform's geospatial reference frame, that the sensor can be properly placed in relation to both its environment and other sensors. This makes it possible to describe a sensor once and then deploy it to many locations, including mobile ones. The implication for on-board data mining, especially for orbital platforms, is large.

Since the physical location of a sensor is not part of its descriptions, one might assume that it would make sense to describe a sensor type, and then have that description applied to a whole group or family of sensors. After all, this is the way XML languages are used to describe groups of data sets and other standard entities. However, each sensor is unique in a variety of ways, requiring a unique XML description. For example, each sensor has unique identifiers such as ID and serial number, as well as additional calibration information specific to that sensor. SensorML is set up to gather this additional information, and can also store a history description that records the changes to the sensor as time progresses.

2.3 Sensor Platforms

A “sensor system” includes one or more sensors and the platform on which they reside. The platform does not make any measurements of physical phenomena, and is thus not a sensor itself. However, the platform does measure its location and orientation, which may be dynamic, as in the case of using an aircraft or a satellite as a sensor platform. The platform’s location and orientation information is very germane to the readings of the sensors aboard that platform. So, while the sensor platform and the sensor itself are considered to be two separate entities, they have a very strong association with one another.

This association is accomplished through the use of “coordinate frames.” The sensor and the platform each have a coordinate frame. The sensor’s coordinate frame locates the sensor on the platform, while the platform’s coordinate frame locates the platform in spatial relation to the larger environment. The coordinate frames are described in terms of coordinate reference systems (CRS). The relationship between a sensor’s CRS and its platform’s CRS is then used to relate both to an external CRS, such as geographic latitude and longitude. These CRSs enable the georegistering of sensors and their measurements. Currently, CRSs are applied only to location, not to temporal information. For example, a relationship between “Earth time” and the delta time of a sensor’s scan start cannot be captured in current CRSs.

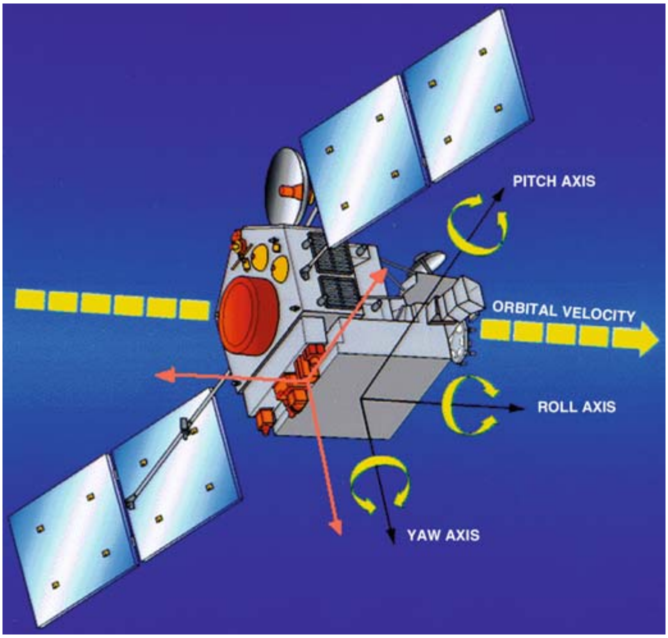

Fig. 3 Relationship of sensor frame (*p*ink) to the moving platform frame (*b*lack) (Courtesy of Open Geospatial Consortium, Inc.)

2.4 Measurements and Response

SensorML can be used to describe properties of a sensor's measurements, specifically, the physical properties being measured by the sensor and the quality of the measurements.

Response characteristics can also be modeled. For example, SensorML can model a sensor's responsiveness to changing environmental conditions. These responses include sensitivity, accuracy, and precision. This information could be very useful to an ensemble data mining approach for dynamically selecting the algorithms most appropriate for the given current conditions.

Overall, by providing a way to describe sensors and system of sensors, and their characteristics and locations, SensorML aids the on-board data mining effort by addressing the problem of georeferencing data from sensors that move.

3 The Use of FPGAs for On-Board Systems

One approach to addressing the problems of limited computing resources and limited power availability is through the use of Field Programmable Gate Arrays (FPGAs) as a means to drive the on-board computing systems. Simply put, FPGAs provide a means to implement software algorithms in hardware, thus increasing speed and hopefully reducing power consumption and heat generation. FPGAs are based on reconfigurable components such as flip-flops and multiplexers, with a bank of SRAM to hold the configuration information and control the functionality of the device. The state of the bits in the SRAM determines the behavior of the device, allowing it to be adapted or customized for different tasks as needed. While somewhat slower than customized hardware, their adaptability makes FPGAs more attractive than a customized chip or hardware solution. The primary downsides to using FPGAs are twofold: the hardware is more expensive, and they can be difficult to program.

Because they are reconfigurable, FPGAs are good tools for prototyping systems. In early FPGAs, the time to reconfigure the system was substantial, limiting their use for actual on-board systems. However, as the cost and time for reconfiguration have dropped (some systems approach reconfiguration times of milliseconds), and as they have become more sophisticated, they have come to be used in production systems. They are especially well suited to data mining applications because they have great potential for parallelism, they perform bit-level operations very efficiently, they can accommodate high bandwidth, and they can be changed dynamically to suit an individual application [68].

There are a number of applications where FPGAs have been used or tested for their applicability to data mining. In one study in Greece, the authors found that implementing the a priori algorithm for data mining using FPGAs was several orders-of-magnitude faster than the fastest available non-FPGA implementation [19]. In other studies, FPGAs were used for image processing and weather modeling, both required substantial numbers of matrix operations. In one case, FPGAs were used in conjunction with a host computer to perform real-time image processing algorithms from an image stream. For each image, line buffering was performed using a 2D digital Finite Impulse Response filter convolution kernel; then noise reduction was performed using a Gaussian 3×3 kernel, followed by edge detection using Prewitt filters. The system performance was substantially faster than a software only solution, but still provided the researchers with the necessary flexibility to adapt the system [15].

In Turin, Italy, FPGAs have been used in a system that collects temperature, humidity, and precipitation data on a routine basis every 15 min and uses that data to make very local weather predictions for the next one, two, and three hours. Such forecasts are especially useful in applications where weather can cause dangerous conditions within a small area. For example, rain can have a substantial impact on the course of a race, and a thunderstorm or fog can shut down an airport. Predicting weather involves considerable data from different sources and is usually computationally intensive. In most cases, observation data is sent to a central computing location, often a large-scale cluster configuration. In this case, however, data mining techniques are employed locally to select the best predictors for a given set of conditions, since the cross-correlations among predictors may change depending on conditions. By implementing such a system using FPGAs, the performance requirements for real-time forecasting were met, and the system was platform independent and adaptable to other applications [14].

In another study using FPGAs for image processing, a single FPGA was used, without any host computer or other co-processors. This created a very inexpensive system with high performance and very low power consumption of only 200 mW, compared to existing systems created to do the same sorts of tasks, which used 600-4,000 mW. The FPGA system was at least comparable in speed to other systems, at a tiny fraction of the cost [16].

The ability to achieve high performance with low expense and power consumption is obviously very appealing for satellites and space applications. One system was designed using FPGAs to ascertain the physical properties of an observed area from satellite data in real time, with the intention of using the system for on-board image analysis and data compression. Data streams from different sensors had to be fused and processed to accomplish this. The system was created by connecting the sensors to a PCI bus, and connecting a host processor and a series of FPGAs to the same bus to process the data. With this simple architecture, the system was faster than Pentium-based PCs by two to three orders-of-magnitude. In addition, the authors suggested that further improvement could be achieved by replacing the bus with interconnections among the FPGAs, since the bus was the primary bottleneck in their studies, and by improving the parallelization of the algorithms used [13].

An FPGA-based system has been proposed for use aboard the Mars scout mission Mars Volcanic Emission and Life (MARVEL), expected to launch in 2011. On-board will be the Mars Atmospheric Trace Molecule Spectroscopy (MATMOS) instrument, observing the sunrise and sunset in a 3-min observation period everyday. These observations will produce far more data than can possibly be downlinked to Earth, but 112 min per day will be available for data processing, which will include heavy use of FFT [17]. That may not sound like much processing time, but considering the environment, it is a substantial commitment of resources.

Several existing systems were considered and discarded because they lacked the processing power to perform this task in the time allotted. However, an FPGA system-on-a-chip solution was found to meet the requirements. The FPGA system is cheap, small, scalable, reconfigurable, and uses very little power. FPGAs are especially suited for use in systems for space missions because they are easier to change, update, and reuse than traditional hardware; they carry less risk of obsolescence; they lend themselves to parallel processing and modularity; and they offer more implementation options than traditional systems. Notably, with an FPGA system, the division of processing between hardware and software is flexible [17].

For this application, an FPGA system was the only available choice that could keep up with the requirements for on-board data processing. The authors estimate that the on-board data processing system they propose will reduce the volume of data for downlink by about 80 times, resulting in an expected savings of tens of millions of dollars over a 2-year mission [17].

Some current space observing missions are using multiple sensors of different types on the same satellite, resulting in large quantities of heterogeneous data to be downlinked and processed. As instruments are deployed farther from Earth, downlinking all of the data becomes less and less feasible, requiring years or decades for deep space missions. For those applications, on-board data fusion becomes critical, and it must be accomplished within the constraints of low power consumption [13]. One clear direction for computing aboard these missions is using FPGAs in a smart payload computer for space, using multiple FPGAs in parallel and large memory banks to hold the data waiting to be processed. Depending on the application, custom co-processors such as a dedicated FFT core may also be used. By implementing these ideas and parallelizing processing as much as possible, new systems can achieve performance that was impossible with traditional processors [17].

By enabling very efficient parallel processing, FPGAs represent a considerable step toward addressing the problems of power, computation, and, to a certain extent, memory and storage in an on-board data mining system. As more sophisticated data mining can be done in an on-board system, the issue of bandwidth will also be alleviated – these systems will be able to summarize, filter, and prioritize data so that the volume of data needing to be downlinked is reduced considerably.

4 Applications for On-Board Data Mining

Although the problems and obstacles associated with on-board data mining are very real, advances in technology have enabled on-board data mining to be performed in some applications. On-board data mining has already proven useful in the areas of autonomous and unmanned vehicles, as well as biometrics.

4.1 Autonomy

One area where on-board data mining has been successfully implemented and deployed is in autonomous spacecraft. For components such as satellites and planetary rovers, the traditional approach has been to gather and store the data on-board the component, and then wait for an opportunity to send it all to ground-based downlink stations on Earth for processing. However, there are problems and severe limitations with this approach, the primary one being that there is usually limited bandwidth available to send the data to Earth [6]. In some cases, a satellite has the capability to collect much more data than it can reasonably transmit. If all data processing and analysis is done at a ground-based facility, any data that exceeds the downlink capacity cannot be used [10]. Also, when interesting events or discoveries are uncovered in the data that comes from a satellite or rover, often further observation of the phenomena would be desirable. If data analysis is performed after the fact, there may be a considerable time lag between the time when a phenomenon is first observed by the instrument and when it is detected on the ground. By the time the data is downlinked and analyzed, retargeting the craft to further study the phenomenon becomes difficult or impossible [6].

However, if the data can be processed on-board the craft, an interesting phenomena can be detected much more quickly, and the craft can react dynamically by retargeting itself for further observations. Furthermore, if data processing occurs on-board the craft, the data can be prioritized and only the interesting or relevant data would be sent to Earth [6]. For instance, if an image is obscured by clouds or other weather that the instrument cannot penetrate, there is no need to waste bandwidth downlinking it. (Of course if scientists are interested in studying clouds, that is another matter.) On the other hand, if an instrument's data contains information about a new and unexpected volcanic eruption, that data should be a high priority for downlinking, and should be able to preempt other lesser data aside in the schedule.

There are several projects in which on-board data mining has been successfully implemented on autonomous space craft. Three of the most successful are EO-1, Mars Odyssey, and rovers on Mars.

4.2 EO-1

Earth Observing-1 (EO-1) is an Earth science oriented satellite, flying in the same orbit as Landsat 7. One of its missions is to test new technologies for remote earth observation [7]. In 2003, the Autonomous Sciencecraft Experiment (ASE) was launched aboard EO-1. The ASE commands EO-1 and retargets the spacecraft in response to events of interest to scientists, capturing further detail of those events as they occur [4]. ASE consists of three parts: on-board science processing algorithms, which analyze data on-board the craft, detect events, and specify an appropriate reaction to those events; on-board planning and scheduling software, which schedules the craft's activities based on analysis of the data, input from the ground, and the craft's capabilities; and robust execution software, which monitors the plan and progress, making sure the craft operates within the bounds of safety even in the face of anomalies [4].

Specifically, science analysis algorithms look for recognizable landscape features or events such as floods, ice formations, clouds, and volcanoes [5]. Before launch, scientists chose locations to monitor events to look for, and an appropriate course of action for each event. They also used archived data to train automatic feature recognition software to recognize specific types of events and landscape features. Once on-board the spacecraft, the algorithms can monitor an area for changes over time or unusual occurrences by comparing multiple images taken over the same location [3].

This is where the strength of data mining in an on-board environment comes into play. As part of the on-board data analysis, scientists developed a Support Vector Machine (SVM) classifier to identify the pixels in an image as water, ice, land, snow, or cloud. Sometimes, if data is missing or bad in some bands, an extra "unknown" class is also used. Knowing the proportion of each of these elements in the image can give clues as to what is happening on the ground. For instance, if an image is not obstructed by clouds and the area has the proper proportions of snow, ice, and water, then that area is flagged as probably containing ice breakup [6]. This type of event can prove useful in studying the effects of global climate change.

When the algorithm detects an event of interest or an anomaly, it can summarize the data, send an alert, change the craft's observation schedule to study the phenomenon further, or simply select only the relevant data to downlink to earth [3].

Once the science analysis algorithms have identified a phenomenon for further study, the on-board planning and scheduling software rebuilds the spacecraft's plan to include further observations, if this is feasible and in keeping with the craft's mission priorities. The Continuous Activity Scheduling Planning Execution and Replanning (CASPER) system is responsible for this task. Based on observations, plans from the ground, and priorities set by the scientists, the CASPER system can retarget the craft for further observations [4]. The CASPER system works by taking all the goals specified for the spacecraft by scientists and by the automated system, considering the current state of the craft, and identifying all conflicts. Then it iteratively modifies the plan until all conflicts are eliminated, using the scientists' specified priorities for data observation [5].

The CASPER system conveys the plan to the robust execution system using spacecraft command language (SCL) [4]. Then the robust execution system executes the plan. It monitors the state of the craft and checks each command for safety before executing it. In the case of a hazard or anomaly, the robust execution system can override the plan generated by CASPER to preserve the safety of the craft [5].

In action, the ASE system operates in a cycle. First, ASE gets prioritized goals from the ground. Then, CASPER makes a plan, choosing the appropriate instruments for each target. The EO-1 spacecraft takes images of the target, and the algorithms analyze those images. If they include changes or phenomena of interest, a new goal of continuing to monitor the location is created. Otherwise, the data in question is marked as not of interest and is not downlinked to earth. If a new goal has been generated, CASPER makes a new plan, and the cycle starts again [5].

The ASE system was created in the face of some difficult challenges. One of the major challenges was communication with earth. The EO-1 spacecraft is in communication with earth for only eight 10-min periods per day. This puts great limitations on the amount of input and control from the ground, and the amount of data that can be downlinked. Meanwhile, the instruments on the craft have limited observation capabilities, so the craft must control itself without being able to see the bigger picture. All this must be done with very limited computing resources. On-board EO-1, ASE was allocated 4 MIPS and only part of the craft's 128 MB of memory. Yet the ASE system has to manage a complex spacecraft with thousands of parts in the hostile environment of space at high financial stakes, ensuring the safety of the craft and maximizing the scientific return from it [4].

Since launch, the ASE system has had considerable impact on the EO-1 mission. The ability for the craft to re-task itself dynamically has enabled much greater flexibility in observations. Previously, plans for the craft were made 5–11 days in advance; with ASE, the timeframe is more like several hours. Also, ASE makes it easier to work around anomalies. Before, the planning and scheduling was done by hand and required considerable thought and knowledge of the spacecraft's capabilities. Automating that task has saved considerable time. Overall, implementing ASE has resulted in cost reductions of over $1 million per year [5].

4.3 Mars Odyssey

As with EO-1, the Mars Odyssey satellite, which observes and maps the surface of Mars, has benefited from on-board data mining efforts. In particular, this craft had the potential to cover more area if there were sufficient bandwidth to downlink the data [10]. Given the success of ASE, scientists wanted to create a similar system to observe Mars [12]. As on EO-1, on-board data mining could enable detection and fast reaction to events as they happen as well as prioritizing the data for downlink. Also, in the course of analysis, additional products such as summaries and alerts are generated with little or no extra effort. These products can be very useful to scientists, and they can be downlinked in place of a whole file when expedient to save bandwidth [11].

In this effort, the Thermal Emission Imaging System (THEMIS) camera on the Odyssey was used. This camera was chosen because of its high spatial resolution and thermal sensitivity [11]. Areas of particular interest to scientists observing Mars include detecting thermal anomalies, dust storms, and the movement of polar ice caps [12].

Detecting thermal anomalies on Mars is of particular interest to scientists because so little is known about Mars. Thermal anomalies are caused by lava flow, frost at low altitude, and very fresh impact craters. They also could be caused by subsurface hydrothermal activity, which would be a significant discovery. An on-board data mining algorithm was developed to detect these thermal anomalies using a threshold based on the season, location, and time of day. If a pixel has a value outside the expected range, it is labeled as an anomaly. Post-processing is conducted to eliminate images with too many pixels flagged, which is indicative of a false alarm [11].

Another area of interest to scientists is estimating aerosol opacity. Specifically, algorithms have been developed using an SVM regression model to flag dust and water ice clouds, which are of interest in understanding the Martian atmosphere [11].

Monitoring the seasonal migration of the polar ice caps on Mars is also of interest to scientists [10]. These ice caps are made of CO_2, and grow and shrink each year with the seasons. This change represents a significant change in the distribution of mass on the planet, which leads to a change in the center of gravity, that is so dramatic it can be observed from earth [10].

In observing the ice caps, one goal is to monitor the location of the edge of the ice cap. To this end, images were taken of the north polar ice cap, on the side of Mars facing the sun, from north to south. Some images contained only ice cap, some contained only noncap land, and some contained the edge. First, the images were classified by whether they contained the edge or not. Since the CO_2 ice has a significantly different temperature than that of nonice cap land, a histogram of the temperatures was used. If a histogram of the temperatures recorded in the image contained two peaks, one represented the temperatures of the ice cap and the other the temperatures of the land, indicating that the edge of the ice cap was contained in the image. If the histogram contained only one peak, the image contained only ice or nonfrozen land [11].

Once the images were classified, the images containing the edge were processed to find the location of the edge. The ice caps do not actually have a distinct edge; rather, the ice gets thinner and thinner, then there is CO_2 frost, then just patches of frost. The scientists declared the point at which less than half the ground had frozen CO_2 to be the edge of the ice cap. The Bimodal Image Temperature (BIT) histogram analysis algorithm was created to find the edge of the ice cap in an image. The algorithm uses the temperature histogram of the image to identify the temperature of the ice cap edge. The minimum point of the dip between the ice cap peak and the land peak represents the temperature of the edge of the ice cap. The algorithm goes through the image and marks each pixel as less than that temperature (ice cap) or greater (nonice land). The northernmost line in the image where less than half of the pixels are marked as ice cap is then declared to be the edge of the ice cap. This location is noted [10].

This algorithm runs in linear time, and it has the advantage of not requiring the data to be calibrated before processing. This is significant, since the data onboard the satellite is not calibrated, and calibrating it would require considerably more processing resources [11]. Another advantage is that once the location of the edge of the ice cap is known, if there is not enough bandwidth to downlink the entirety of the images containing it, just the location data can be transmitted to earth [10].

One unexpected result of these efforts is the discovery of an additional band of water ice sometimes present at the edge of the polar ice cap. This band appears as

Fig. 4 Mars Odyssey's mission is to map chemical elements and minerals on the surface of Mars, look for water in the shallow subsurface, and analyze the radiation environment to determine its potential effects on human health (Image Courtesy of NASA)

a third peak in the temperature histogram of some images, too warm to be CO_2 ice but too cold to be nonice land. It grows wider as the CO_2 ice recedes in the spring. Prior to this work, its existence was only guessed at [11].

As with running ASE on EO-1, this effort faced major constraints. The processor aboard the Odyssey runs at 20 MHz, and only of 20% capacity, and 40 MB of memory was allocated for on-board data analysis. With those processing resources, the algorithm had to be fast enough to keep up with the data collection, which occurs at 9,600 pixels per second [11].

4.4 Mars Rover

Like the previous satellite projects, the Mars rovers face constraints of limited downlink capacity and major delays, making full control from Earth nearly impossible. The traditional approach of manually selecting targets for sampling based on the previous (Martian) day's images, or guessing and blindly directing the rover toward suspected targets, was fraught with delays and inaccuracies [8]. Meanwhile, as rover guidance and distance measurement systems improved, the rovers were able to travel farther, gathering more data, but with no increase in bandwidth to downlink that data [9].

Fig. 5 The Mars Rover mission has been wildly successful in exploring the Martian surface (Image Courtesy of NASA)

The On-board Autonomous Science Investigation System (OASIS) was developed to allow each rover to analyze data on-board, identifying items of interest in its images and retargeting itself as necessary for further study. OASIS manages geologic data on-board the rover, identifies terrain features and selects targets of interest, and plans and schedules rover movement using the same CASPER system used on EO-1 [9].

The Mars rover mainly seeks to identify different types of rocks in the landscape. To do this, feature detection is used. First, an image is split into ground vs. sky. If the sky is in the image, edge detection is performed to find the horizon, seed areas of low variance are selected, and the sky is identified by filling in from those seeds [9].

Once sky is identified, clouds are detected by finding areas of high variance in the sky. For rock detecting, the sky is masked out. The remaining image is normalized, smoothed, and the edges are detected. Rocks are identified by using an edge walker to find closed shapes in the image [9].

Once the rocks in the image have been identified, the algorithm picks specific points on individual rocks as targets [8]. Rock properties such as albedo, shape, size, and texture are estimated and used to prioritize the rocks for further study, since these characteristics indicate the type of rock. Unusual rocks are marked with a higher priority, but care is taken to make sure representative rocks are also included [9].

Once this processing has occurred, the algorithm can flag images as useful or not. For example, images with no clouds are of no use to scientists who are studying clouds and should not be sent to earth unless they contain something else of interest. Once the targets have been identified, new goals are generated and sent to CASPER, which creates a new plan for the rover, studying as many targets as possible according to priority [9].

Future expansions to this research include more sophisticated rock property gauging, recognizing more features, being able to recognize the same target from different angles and distances, and being able to identify larger geographical boundaries such as hills, plains, and river channels. However, even without those improvements, this effort has benefited the rover program by enabling a rover to identify new targets on the fly, examine interesting things as it discovers them if resources are available, and prioritize data for downlink to earth, thus maximizing scientific benefit from the resources available [9]. Clearly, adding on-board data mining capabilities to deep space missions is the next step.

4.5 Deep Space

As space exploration moves farther from Earth, the large latencies, high error rates, and limited communication opportunities involved mean more potential benefit from performing data mining on-board the instrument or spacecraft [35]. Future deep space missions will include autonomous operations, including on-board fault-detection and repair as well as on-board data mining [36]. Intelligent scheduling, data prioritizing, and on-board data mining will be crucial to the maximization of resources in space [35].

Autonomy is already being pursued in deep space exploration. In 1998, Deep Space 1 (DS1) was launched for the purpose of testing the viability of new, high-risk technologies. It was the first deep space probe to use an autonomous on-board navigation system, called AutoNav. It was given a baseline trajectory, plus a set of information about the locations of asteroids, planets, stars, and the intended targets for DS-1. As DS-1 separated from the launch vehicle, AutoNav worked with the Attitude Control System (ACS) to determine the craft's attitude based on the positions of the stars, and move the craft as needed. This included turning the spacecraft to be in optimal position to catch the sun for power generation [35].

To perform these navigation functions, the AutoNav system relies on images gathered from on-board cameras. The AutoNav system plans which upcoming images it requires, and works with the ACS system to position the craft so that those images can be captured. The images are processed, and the necessary information is extracted to compute the current location and attitude of the craft. The AutoNav system is then able to calculate the track from the current position to the next and make an assessment of the craft's progress. If necessary, AutoNav works with ACS to perform a course correction [38].

In testing, the AutoNav system was able to get the spacecraft within 2.5 km of its target locations and arrive within 5 s of its scheduled times – which is fairly impressive for an autonomous system with all of space to run around in. This is all the more impressive because there were considerable problems with the CCD camera on which the system relied. Furthermore, this type of navigation combined with continuous low-thrust propulsion had never been attempted before, either automatically or manually [38]. After some ground-based parameter tuning, AutoNav's accuracy was further improved [40].

After the initial tests ended, the AutoNav system continued to control DS1, including maneuvering the craft for an encounter with comet Borrelly. Again, AutoNav performed successfully, maneuvering the craft within 2,200 km of the comet despite the failure of another major component of the ACS, the Stellar Reference Unit (SRU) [39]. As a result of its navigational success, DS1 was able to send back the best images and science data ever collected from a comet [41].

DS1 also used an autonomous remote agent to plan and update the spacecraft's schedule based on the mission goals and the state of the craft. This system performed well in tests, including working around simulated faults [40], proving the viability of such autonomous navigation capabilities.

By incorporating these autonomous systems into DS1, NASA was able to reduce costs considerably. The ground operations team averaged only 50 full-time equivalent personnel, resulting in considerable cost savings over the larger team that would be required for a more hands-on mission [40]. The total cost for the DS1 mission was under $150 million [41] – relatively inexpensive in the world of space craft.

As an additional note, autonomous spacecraft can also be supported by autonomous ground stations. The Deep Space Terminal (DS-T) was created as a terminal in the Deep Space Network (DSN) to monitor instruments in space without human intervention. In 1998, DS-T successfully demonstrated its intended function by autonomously receiving, processing, recording, and distributing data from a probe orbiting Mars to the rest of the network [37].

5 Unmanned Vehicles

Like autonomous vehicles in space, unmanned vehicles, whether on the ground (UGVs), in the air (UAVs), or under water (AUVs), represent a great opportunity to take advantage of on-board data mining. For these vehicles to make the leap from remotely controlled devices to autonomous objects, they need to be able to perform their own planning, scheduling, and control tasks. One approach for accomplishing this, suggested by Tim Grant [46], follows the Observe–Orient-Decide–Act model (sometimes referred to as the OODA-Loop). The OODA-Loop was originally developed for use by fighter pilots to quickly make decisions and act on those decisions [46] and has been used successfully in a number of AI-based systems. In addition, complete path planning algorithms have been developed to find optimal paths using two different approaches: probabilistic roadmaps (RPM) and rapidly exploring random trees (RRT). Since, in the case of some UAVs, the vehicles in question would be flying in the same airspace as commercial airliners, collision avoidance is crucial. The systems also need to be able to replan based on new information, such as obstacles or updated no-fly zones. These approaches performed well in a test flight involving fully deployed UAVs [48]. Similarly, navigation systems for unmanned ground vehicles (UGVs) have been proposed, using GPS and compass readings for location and orientation [49].

For small devices, movement similar to that of birds or insects is often useful. This behavior is called swarming. Basically, each vehicle or sensor moves based on simple rules, but as they interact, an emergent intelligent behavior seems to occur, as seen with ants looking for food – each ant's behavior may seem random, but taken together the collective hive behaves in an intelligent manner. As such algorithms are optimized for parallel processing, they can be very useful for determining UAV behavior, especially when there are many small, unsophisticated UAVs working together [45].

Another approach is to use fuzzy logic for planning and control of UAVs. By employing fuzzy decision trees, a plan can be developed that accounts for factors including risk to the vehicle, fuel consumption, other costs, and mission goals and priorities, while determining the optimal trajectory for each UAV. Using fuzzy decision trees, the UAVs can also collaborate automatically without human intervention, again taking into account each vehicle's goals, safety, and priorities [43].

UAVs are becoming increasingly important to the military, especially for surveillance missions. In particular, urban surveillance presents challenges of seeing around corners and picking out the relevant information from a busy image. One application in which on-board data mining can be particularly advantageous is in target tracking. In this case, a UAV can go into a dangerous environment and pursue a target such as a suspected terrorist or a vehicle. To accomplish this, the UAV needs to be able to identify the target from the images it collects, track its movement though a series of temporally separate images, and navigate to follow the same path. This is not an easy task, but must be performed on-board. Bandwidth constraints limit the ability to transmit the images for processing elsewhere and to send feedback to the vehicle – the delays involved would be unacceptable. This is one application where the advances in on-board data mining result in clear, immediate benefits to the safety of humans – expendable and autonomous UAVs are sent into harm's way and convey results back to humans located in relative safety [50].

In one approach addressing this type of application, Support Vector Regression was used to establish the location of a stationary enemy radar station. This was done by tracking the radar pulse as received by three UAVs flying together in a triangular formation and using that information to calculate the location of the source of the signal. This work represents a first step toward a scalable solution that would track stationary and moving targets using multiple mobile platforms [51].

UAVs are also coming into wider use by civilian agencies, including the department of transportation. They provide a lower cost alternative to manned vehicles while allowing faster and more flexible deployment compared to fixed cameras mounted along highways. For remote areas, UAVs may be the only cost-effective option for traffic monitoring. Beyond simply transmitting video of traffic flow to ground sites for processing, UAVs with on-board data mining capabilities have the potential to recognize accidents or emergencies and alert emergency personnel with images and location information [44].

There are many opportunities for applying on-board data mining in underwater applications as well. In one project, autonomous underwater vehicles (AUVs) were used to study the sea floor, looking for bacterial mats living near offshore mud volcanoes or areas rich in organic matter. As multiple AUVs worked together, their images could be combined to produce video mosaics and georeferenced images showing the areas explored. Scientists on one such project developed the IBU software to automatically analyze the images, a process that had been completely manual before [52].

As this project produced large volumes of data, thousands of images per campaign, on-board processing was used to distinguish the relevant parts of an image from the background and send only the relevant parts on for further processing on the surface. After performing manual and automated analysis of a set of 2,840 video images and comparing the automated results with the manually identified images, scientists found that the automated analysis software exhibited better than 90% precision. These positive results can be extended to move more of the processing to the UAVs, allowing them to become more autonomous and make decisions based on their analysis of images as they are captured [52].

New unmanned underwater vehicles are being developed that can autonomously control themselves and collaborate with each other while floating freely with the ocean currents. By moving with the currents rather than being stationary or self-propelled, these vehicles can observe the ocean life in its natural state without disturbing it. This sort of device is also uniquely suited to tracking such things as oil spill spread, pollution dispersion, and plankton community evolution. In one test case, individual prototype devices have been deployed successfully. The eventual goal is to release many of these devices together and allow them to interact through an acoustic network. Since these devices are designed to travel without human control and the range of underwater modems is only a few miles, bandwidth will be at a premium [47]. Any on-board data mining that these devices can do will be very valuable in reducing the communication requirements.

6 Biometrics

In addition to enabling instruments in hostile or distant environments to perform autonomously, on-board data mining technology is also enabling advances in biometric systems. Biometric systems are one or more sensors used to monitor some facet of the human body. Such sensors may be placed in well connected areas – for example, the fingerprint systems now in use at many airports. But others either are placed in a covert way to surreptitiously monitor a location or are portable systems meant for use by individuals to monitor their health. Thus, many of the usual problems and approaches to on-board mining come into play with these systems.

As an example, computer-based authentication systems are commonplace – key cards, username/password systems, etc. Many such systems currently use knowledge-based or token-based authentication: authentication based on something you know (such as a password or PIN) or something you have (such as a card or a key). However, these systems are prone to unauthorized access if someone obtains the password or steals the card or key. Biometrics offers the possibility for unique and secure identification of each and every person. There are two primary types of biometric data: static, involving a physical attribute such as fingerprints or an iris scan; and dynamic, involving behavior, such as keystroke dynamics, signature dynamics, or even a person's gait [21].

To date, most single-source biometric systems have relatively poor performance; however, if multiple sources of biometric data are combined, biometric systems can approach near-flawless performance. For example, a system may check the user's fingerprint, gait, and voice for identification. By using more than one attribute, the user may still be authenticated if the reading of one attribute is corrupted in some way. For example, if the user has a cold, voice recognition may be impacted. Similarly, dust on the lens or unfavorable lighting may interfere with facial recognition. Using multiple attributes allows a positive identification to be made despite such problems. However, it results in a high volume of heterogeneous data to process. One method that has been found effective in such a case is Bayesian Model Averaging with Decision Trees and a sweeping strategy. This method is particularly effective in situations where risk evaluation is crucial, such as at border checkpoints and transportation hubs [21].

Even when using just one biometric attribute, data volume can still be an issue, and identification is subject to problems such as lighting variation. For example, although the use of fingerprints for identification is one of the most mature and proven biometrics, fingerprint recognition and classification are still a difficult problem. To improve accuracy and reduce processing time, one approach is to normalize the images, extract features, classify the image, and then recognize the fingerprint based on its classification. Normalization includes adjusting all images to the same darkness/brightness, and adjusting the orientation. Images can also be enhanced by thinning and increasing contrast. Then a line detector can be used for feature extraction of uniquely identifying characteristics, in this case detection of ridges and furrows. A given pattern of ridges and furrows is classified as a whorl, arch, tent, etc., using a neural network approach. Finally, using a crisp k nearest neighbor algorithm, the image is matched against selected images in the database based on the classification. This saves considerable processing time in not having to search the entire database. In tests, this approach achieved a 97.4% recognition rate [22].

Similarly, face recognition is subject to difficulties resulting from cropping, differences in pose angle, and differences in lighting. Several approaches have been used, including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and eigenfaces, an approach using feature vectors and the Euclidean distance between them. However, all of these methods are very sensitive to outliers in the data, specifically "noisy" images with lighting, cropping, or pose angle different from the images in the database, making their use in convert situations difficult. However, one study found that by implementing an automated system to filter out images whose feature vectors represent outliers, that is, noisy images, the recognition rate of a face recognition system could be improved by 10–20% [26].

These incremental types of improvements become crucial in on-board applications, where processing time, power, and network connectivity may all be limited. For example, in border control systems, the person must be identified accurately and processed expediently. Considerable discussion has occurred regarding including biometric data in passports. In 2002, the International Civil Aviation Organization (ICAO) put forth standards for future passports, including the use of smart cards for storing biometric data [31]. The benefits of more reliable identification for border control and security purposes are obvious. However, this plan is not without obstacles. To develop and maintain a database of biometric data, fingerprints and faces for example, for travelers worldwide is a large and costly endeavor. Identifying a traveler from such a large database in a timely fashion is another problem. Also, the cost of issuing passports with secure smart card technology must be considered, as well as the security of the data once in the smart card and in the database, and user resistance to giving biometric data. The security of the system is of utmost importance: if the biometrics data can be stolen or manipulated, the entire effort may do more harm than good. In 2005, an early Dutch biometric passport was cracked in 2 h using a PC, so this is a real issue [29]. However, some progress has been made. Visitors to the U.S. from countries participating in a visa waiver program have fingerprint scans and a digital photograph collected at the time a visa is issued. At that time, these biometrics are compared against a database of criminals and terrorist suspects. If the visitor is not among the banned, the visa is issued. The biometric data is then used to verify the visitor's identity on entry and exit into the US. [32] [33].

In the commercial world, there are a number of examples of biometrics systems that have been deployed successfully. IBM Thinkpads and Korean LG mobile phones have both incorporated fingerprint recognition for user authentication. Also, in a pilot study, the Port of Palm Beach, Florida, used photographs and fingerprints to identify and track visitors, and to keep out banned visitors. Furthermore, the Pinellas County Sheriff's Office in Florida implemented a face recognition system using digital cameras and docking stations in patrol cars. In less than 1 year, 37 identifications were made leading to arrests. None of these identifications would have been made without the on-board systems in the patrol cars [31].

A more unusual application for on-board data mining with biometrics is used for mood analysis. The Mood Phone analyzes acoustic information in speech and ascertains the speaker's mood. This information is conveyed in real time to the listener by a color coded light that also varies in intensity based on the intensity of the speaker's mood. This information can be very useful to people who are impaired in detecting others' moods themselves, including people with autism or Asperger's syndrome. Another emotion detecting application uses a glove with sensors to monitor skin temperature, conductivity, and heart rate. This system can estimate the user's emotional state with up to 75% accuracy. One possible application for this technology is for use with students, to detect when they become frustrated with a training scenario and need personal attention [23].

Biometrics such as command use frequency and habits based on time of day can be used to detect intrusion and thwart an attack to a computer network. One study used these factors to achieve a 90% recognition rate for users of a university system for users who had executed at least 50 commands [24]. Haptics, including pressure and torque when using a stylus and feedback-generating pad, have also been shown to uniquely identify users [25]. Even simpler attributes such as typing style and idle time patterns can be used to identify a user. For example, in a system developed for online gaming to stop account theft and sharing one account among several people, the patterns of idle time between commands were found to uniquely identify users. In this study, activity logs for 287 users of the game Angel's Love were analyzed. Any users who logged fewer than 200 min during the study time were eliminated. For the remaining users, there was a high correlation between active time (intervals where the character moved continually, with no pauses of a second or more) and idle time (intervals of 1 s to 30 min when the character did not move). The idle time distribution had greater variation among users, varying in central tendency and overall shape, whereas the active time varied considerably less among users, so the idle time was used as the indicator. In this study, when the length of playing history was fixed to 200 min, user detection could be performed in 20 min with over 90% accuracy. The addition of other factors, such as movement patterns in the game and mouse/keyboard behavior, could speed up the detection process [28].

Keyboard behavior, specifically typing style, has been shown to be effective in user authentication and requires no special hardware [30]. Using the delays between pairs of keystrokes, users can be identified uniquely when entering a user ID and password [20, 30]. In one study, only eight rules were required to identify the legitimate users with over 95% accuracy. Specifically, the first few and last pairs of characters were enough to make a correct classification – the entire user ID and password do not need to be stored or analyzed. Also, the legitimate user tends to type his or her login ID faster than any other person [20].

On-board data mining has life-saving potential when applied to health monitoring and alert systems. Sets of sensors that are worn continually allow for a better understanding of the patient's baseline state as well as immediate notification in the event of an anomaly such as a heart attack or stroke. Traditional in-home health monitoring systems were awkward and cumbersome, relying heavily on wires for communication. They had short memory banks of 24 h or less and no processing capabilities on their own. Instead, systems are being developed that can store and process several weeks of data while withstanding the conditions of ordinary life, including temperature extremes, vigorous activity, and sleep. The sensors are smaller, less obtrusive, noninvasive, and nonirritating to the skin. By combining these sensors with intelligent processing and on-board data mining software, a significant improvement is realized. Health deterioration can be detected earlier and health care providers notified, while reducing the need for in-person medical appointments and monitoring. Furthermore, these advanced systems can provide a valuable bank of detailed information about the patient's condition to the medical practitioners, much more so than with traditional monitoring systems or occasional office visits [55].

Systems have been designed to accomplish this task. In one proposed system, a Cyberjacket, a modular wearable computing platform, is augmented with an ECG, and oxygen saturation monitor, and a temperature sensor. The Cyberjacket already includes a GPS device and accelerometers to detect location and motion.

By combining these components, a reasonably unobtrusive and mobile system can be created, enabling health conditions to be monitored in the patient's everyday life and referenced with context information including time of day and location [54].

A similar system has been proposed using a Smart Shirt with sensors for heart rate, body temperature, and blood pressure, as well as a GPS device to record location. Data from the sensors would be fed into a PDA using Bluetooth. Then the PDA would display signal data and perform quick analysis on the data. In the event of an anomaly or emergency, the PDA would alert medical personnel, using location information provided by the GPS to direct them to the patient. Under normal circumstances, the PDA would also feed the data to a larger server, which would perform further data mining using association rules and classification [53].

A third system, the Wearable Wireless Body/Personal Area Network (WWBAN) has been implemented and tested. This system involves a three-tier network: the wearable sensors send data to a PDA using a short-range wireless network, the PDA passes the data on to a home or central server using wireless LAN protocols, and then the home server can periodically update the patient's record at the hospital or doctor's office using a secure internet connection. The PDA includes personal server (PS) software, providing an interface to display the data to the user while monitoring the user's health and providing feedback, including alerts in the event of an anomalous condition. Using data from the motion sensors, the PS can discriminate the general activity of the user (sitting, walking, running, lying down) and incorporate that information with the health metrics provided by the sensors and the patient's history to give a more complete picture of the patient's condition and status [57,60].

These systems and others like them have great potential for many health applications in addition to ongoing monitoring of vital signs. Another system, called LiveNet, uses classifiers to distinguish among activities such as walking, sitting, and biking, as well as more subtle activities such as head-nodding or shaking. The LiveNet system also incorporates voice processing to detect stress and emotional arousal. In addition, the system can detect shivering with up to 95% accuracy using real-time classifier systems based on Gaussian Mixture Models. One application for this is in monitoring soldiers working in harsh climates and determining their hypothermia risk. LiveNet is also in pilot testing for monitoring Parkinson's disease symptoms, epilepsy seizures, and even depression. The continual monitoring and analysis provided by the system has great potential in helping doctors tailor treatments based on individual patients' true symptoms and experiences, rather than heuristics and average dosages [58].

Data from wearable sensors can also provide insight into stroke patients' recovery and rehabilitation. By using wearable sensors to monitor stroke patients' activities, therapy can be better tailored to their needs. During a test of the system, patients performed tasks including lifting a pencil, turning a key, flipping a card, and other tasks often used to assess a stroke patient's level of impairment. Using algorithms from the data mining toolkit Waikato Environment for Knowledge Analysis (WEKA) and linear regression models, features from each task were analyzed, and models were built to predict the patient's scores on clinical tests, getting within 10% of the average clinical score [59].

7 Sensor Networks

As the field of on-board data mining advances, integrating the data from multiple sensors and getting them to communicate and work together is the next step. Deployment of multiple networked sensors is becoming a more and more common method to monitor large geographic areas. This is due in part to the declining costs of such sensors, but also to the improvements in both wired and wireless network connections. Such networking is the cornerstone of sensor networks. In order for the system to derive the maximum advantage from the gathered information, the components in the network must be able to communicate their observations to each other.

NASA has been at the forefront of such networks for some time. However, it is the military and homeland security interests that are increasingly pushing the technologies in this area. For example, the Defense Intelligence Agency's 5-year plan for leveraging basic research describes a sensor network consisting of intelligent and autonomous space-borne, airborne, and ground-based sensors [61]. These sensors will act independently of one another, yet each will be capable of both publishing and receiving sensor information, observations, and alerts among other sensors in the network. Furthermore, these sensors will be capable of acting on alerts and information received from other sensors in the network, perhaps altering acquisition properties of their instruments, changing the location of their platform, or updating processing strategies for their own observations to provide responsive information or additional alerts.

Such autonomous and intelligent sensor networking capabilities provide significant benefits for collections of heterogeneous sensors within any environment, including those used for treaty verification, covert monitoring, and changing battlespaces, but they are crucial for multi-sensor observations and surveillance. For these applications, real-time communication with external components and users may be inhibited, and the environment may be hostile.

In all environments, mission automation and communication capabilities among disparate sensors will enable quicker response to interesting, rare, or unexpected events, especially important in time-critical situations such as chemical detection, missile firings, or battlefield situations. Another advantage of an intelligent network of heterogeneous sensors is that all of the sensors can benefit from the unique capabilities of each sensor in the network. There are a number of efforts underway to manage real-time sensor data. Some areas of focus include data compression and reduction, sensor fusion, real-time data mining, and other operational level algorithms (e.g. [62–66]). Many of these approaches deal quite effectively with streaming data in resource-constrained environments such as sensor networks.

NASA and the National Science Foundation have a number of sensor network efforts currently in operation and a number that are in the research and development phases. Some of these are aimed at space-based applications, but many are ground-based systems geared toward environmental, weather, biological, and health monitoring. For example, several sensor networks have been deployed to monitor volcanic activity around the world, including some on the US mainland as well as in Hawaii and Alaska. Most of these systems are comprised of sensors deployed around the perimeter of known active volcanoes, and are diligently measuring for any change in activity (e.g. [67]). Some of these networks are also linked with satellite data such as NASA's Moderate Resolution Imaging Spectroradiometer (MODIS), which takes thermal measurements on a global basis. For the people living and working near such geologic features, any change that may signal an imminent eruption is obviously of great interest. Most of these networks include some level of data processing within the network itself, including limited amounts of data mining.

Another area of intense research is in the monitoring of coastal areas. Applications include border protection efforts, using networks of cameras to monitor ports, and even tsunami warning systems that are being deployed after the devastating Indian Ocean tsunami in 2004. The use of sensor networks for environmental monitoring is becoming more widespread as well. All of these applications represent areas where on-board data mining provides added value by enabling information gathering that was previously infeasible or by supporting the optimized use of resources.

8 Conclusion

On-board data mining represents a powerful technology with the potential to improve the efficacy and efficiency of a range of applications. Great strides have already been made in the areas of biometrics, autonomous and unmanned vehicles, and other areas, resulting in new exploration, better data availability to scientists, cost savings, and improvements in security and medicine. As technology improves, harnessing the power of on-board data mining will become increasingly feasible. It is likely that you will be seeing its use in applications in your own work and even your own home and vehicles. This will be a field to watch for exciting new developments.

References

1. Organization web site including all standards documents, http://www.opengeospatial.org/ogc
2. SensorML, http://vast.uah.edu/SensorML/
3. S. Chien, T. Debban, C. Yen, R. Sherwood, R. Castano, B. Cichy, A. Davies, M. Burl, A. Fukunaga, Revolutionary deep space science missions enabled by onboard autonomy. International Symposium on AI, Robotics, and Automation in Space, 2003
4. D. Tran, S. Chien, R. Sherwood, R. Castano, B. Cichy, A. Davies, G. Rabideau, The autonomous sciencecraft experiment onboard the EO-1 spacecraft. AAMAS, 2004
5. G. Rabideau, D. Tran, S. Chien, B. Cichy, R. Sherwood, Mission operations of earth observing-1 with onboard autonomy. Proceedings 2nd IEEE International Conference on Space Mission Challenges for Information Technology, 2006

6. R. Castano, N. Tang, T. Doggett, S. Chien, D. Mazzoni, R. Greeley, B. Cichy, A. Davies, Onboard classifiers for science event detection on a remote sensing spacecraft. Proc. 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006, pp. 845–851
7. NASA: Earth Observing-1, http://eo1.gsfc.nasa.gov/new/general/index.html. Accessed 1 Feb 2008
8. R. Castano, T. Estlin, D. Gaines, A. Castano, B. Bornstein, C. Chouinard, R.C. Anderson, M. Judd, Automated target selection for opportunistic rover science. 37th Lunar and Planetary Science Conference, 2006
9. R. Castano, T. Estlin, R. Anderson, D. Gaines, A. Castano, OASIS: onboard autonomous science investigation system for opportunistic rover science. J. field robot. **24**(5), 379–397 (2007)
10. K. Wagstaff, R. Castano, S. Chien, A. Ivanov, E. Pounders, T. Titus, An onboard data analysis method to track the seasonal polar caps on Mars. International Symposium on AI, Robots, and Automation in Space, 2005
11. R. Castano, K. Wagstaff, S. Chien, T. Stough, B. Tang, On-board analysis of uncalibrated data for a spacecraft at Mars. 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2007, pp. 922–930
12. K. Wagstaff, J. Bandfield, R. Castano, S. Chien, M. Smith, Dust storms and water ice clouds: feature detection for use onboard THEMIS. 37th Lunar and Planetary Science Conference, 2006
13. Y. Cai, Y. Hu, Onboard inverse physics from sensor web. 2nd IEEE International Conference on Space Mission Challenges for Information Technology, 2006
14. T. Meindl, W. Moniaci, D. Gallesio, E. Pasero, Embedded Hardware architecture for statistical rain forecast. Res. Microelectron. Electron. **1**, 133–136 (2005)
15. J.A. Kalomiros, J. Lygouras, Design and evaluation of a hardware/software FPGA-based system for fast image processing. Microprocess. Microsyst. (2008). doi:10.1016/j.micpro.2007.09.001
16. M. Lorenz, L. Mengibar, E. SanMillan, L. Entrena, Low power data processing system with self-reconfigurable architecture. J. Syst. Arch. **53**(9), 568–576 (2007)
17. P. Pingree, J.-F. Blavier, G. Toon, D. Bekker, An FPGA/SoC approach to on-board data processing enabling new mars science with smart payloads. IEEE Aerospace Conference, 2007, pp. 1–12
18. Z. Baker, V. Prasanna, An architecture for efficient hardware data mining using reconfigurable computing systems. 14th Annual IEEE Symposium on Field-Programmable Custom Computing Machines, 2006, pp. 67–75
19. Z. Baker, V. Prasanna, Efficient hardware data mining with the apriori algorithm on FPGAs. 13th, Annual IEEE Symposium on Field-Programmable Custom Computing Machines, 2005, pp. 3–12
20. K. Revett, S.T. de Magalhaes, H. Santos, Data mining a keystroke dynamics based biometrics database using rough sets. Portuguese conference on Artificial Intelligence, 2005, pp. 188–191
21. C. Maple, V. Schetinin, Using a Bayesian averaging model for estimating the reliability of decisions in multimodal biometrics. First International Conference on Availability, Reliability, and Security, 2006
22. K. Umamaheswari, S. Sumathi, S.N. Sivanandam, K.K.N. Anburajan, Efficient finger print image classification and recognition using neural network data mining. International Conference on Signal Processing, Communications, and Networking, 2007, pp. 426–432
23. J. Krikke, B. Alfonsi. “In the news.” IEEE Intelligent Systems, vol. 21, issue 3, Jan.-Feb. 2006, pp. 102–104
24. H.Z. Bing, V.P. Shirochin , S. Jun, An intelligent lightweight intrusion detection system (IDS). IEEE Region 10 TENCON, 2005, pp. 1–7
25. R. Iglesias, M. Orozco, J. Valdes, A. El Saddik, Behavioral features for different haptic-based biometric tasks. IEEE International Workshop on Haptic, Audio, and Visual Environments and Games, 2007, pp. 102–106

26. S. Berrani, C. Garcia, On the impact of outliers on high-dimensional data analysis methods for face recognition. Proc. 2nd International Workshop on Computer Vision Meets Databases, 2005, vol. 160, pp. 43–49
27. R.B. Rao, J. Bi, G. Fung, M. Salganicoff, N. Obuchowski, D. Naidich, LungCAD: a clinically approved, machine learning system for lung cancer detection. Proc. 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2007, pp. 1033–1037
28. K. Chen, L. Hong, User identification based on game-play activity patterns. Proceedings of 6th ACM SIGCOMM Workshop on Network and System Support for Games, 2007, pp. 7–12
29. T. Kwon, H. Moon, Biometric authentication for border control applications. Trans. Knowl. Data Eng. doi:10.1109/TKDE.2007.190716
30. A. Guven, I. Sogukpinar, Understanding users' keystroke patterns for computer access security. Comput. Secur. **22**(8), 695–706 (2003)
31. Biometrics enter mobile world, Biomet. Technol. Today **13**(8), 10–11 (2005)
32. U.S. Department of Homeland Security, US-VISIT: How it works. Accessed 22 Feb 2008
33. U.S. Department of Homeland Security. US-VISIT biometric exit process to improve national security. Accessed 22 Feb 2008
34. T.A. Carper, M.D. Gardiner, Biometric Embedded Device. U.S. Patent 20080040615, filed 2007
35. R. Slywczak, Developing autonomous missions through intelligent on-board architectures. Proceedings of IEEE Conference on Networking, Sensing and Control, 2005, pp. 359–364
36. L. Alkalai, Perspectives on dependable computing for solar system exploration. 2002 Pacific Rim International Symposium on Dependable Computing, 2002
37. L. Paal, N. Golshan, F. Fisher, E. Law, W. Veruttipong, M. Stockett' Deep space terminal demonstration. The Telecommunications and Mission Operations Progress Report, TMO PR 42-138, April-June 1999
38. J.E. Reidel, S. Bhaskaran, S. Desai, D. Han, B. Kennedy, T. McElrath, G.W. Null, M. Ryne, S.P. Synnott, T.C. Want, R.A. Werner, Using autonomous navigation for interplanetary missions: The validation of deep space 1 AutoNav. International Conference on Low-Cost Planetary Missions, 2000
39. S. Collins, Deep space 1 flight experience: adventures on an ion drive. 25th Annual AAS Guidance and Control Conference, 2002
40. M.D. Rayman, P. Varghese, D.H. Lehman, L.L. Livesay, Results from the deep space 1 technology validation mission. Acta Astronaut. **47**(2–9) , 475–487 (2000)
41. NASA, Deep Space 1, http://nmp.jpl.nasa.gov/ds1/. Accessed 3 Mar 2007
42. NASA, Deep Space 1: Quick Facts. Accessed 3 March 2007
43. J.F. Smith III, T.H. Nguyen, Fuzzy decision trees for planning and autonomous control of a coordinated team of UAVs. Proceedings of SPIE 6567, 656708, 2007
44. S. Srinivasan, H. Latchman, J. Shea, T. Wong, J. McNair, Airborne traffic surveillance systems: video surveillance of highway traffic. In Proceedings of the ACM 2nd International Workshop on Video Surveillance & Amp; Sensor Networks, New York, NY, USA, October 15, 2004. VSSN '04 (ACM, New York, 2004), pp. 131–135
45. J.J. Corner, G.B. Lamont, Parallel simulation of UAV swarm scenarios. In Proceedings of the 36th Conference on Winter Simulation, Washington, DC, December 05 – 08, 2004. Winter Simulation Conference. Winter Simulation Conference, pp. 355–363, 2004
46. T. Grant, 2005. Unifying planning and control using an OODA-based architecture. In Proceedings of the 2005 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries (White River, South Africa, September 20–22, 2005). ACM International Conference Proceeding Series, vol. 150. South African Institute for Computer Scientists and Information Technologists, pp. 159–170

47. J. Jaffe, C. Schurgers, Sensor networks of freely drifting autonomous underwater explorers. In Proceedings of the 1st ACM International Workshop on Underwater Networks, Los Angeles, CA, USA, Sept 25–25, 2006. WUWNet '06. ACM, (New York, NY, 2006), pp. 93–96
48. M. Wzorek, P. Doherty, Reconfigurable path planning for an autonomous unmanned aerial vehicle. Hybrid Information Technology, 2006. ICHIT'06. Vol 2. International Conference on. Volume 2, Nov. 2006, pp. 242–249
49. B.-J. Yoon, M.-W. Park, J.-H. Kim, UGV(unmanned ground vehicle) navigation method using GPS and compass. SICE-ICASE, 2006. International Joint Conference, Oct. 2006, pp. 3621–3625
50. B. Ludington, J. Reimann, G. Vachtsevanos, I. Barlas, Target tracking with unmanned aerial vehicles: From single to swarm vehicle autonomy and intelligence. Control and automation, 2006. MED '06. 14th Mediterranean Conference, June 2006, pp.1–6
51. B. Sundaram, M. Palaniswami, S. Reddy, M. Sinickas, Radar localization with multiple unmanned aerial vehicles using support vector regression. Intelligent sensing and information processing, 2005. ICISIP 2005. Third International Conference , 14-17 Dec 2005, pp. 232–237
52. K. Jerosch, A. Ldtke, M. Schlter, G.T. Ioannidis, Automatic content-based analysis of georeferenced image data: Detection of Beggiatoa mats in seafloor video mosaics from the Hkon Mosby Mud Volcano. Comput. Geosci. **33**(2), 202–218 (2007)
53. P.-T. Cheng, L.-M. Tsai, L.-W. Lu, D.-L. Yang, The design of PDA-based biomedical data processing and analysis for intelligent wearable health monitoring systems. Computer and Information Technology, 2004. CIT '04. The Fourth International Conference, 14–16 Sept 2004, pp. 879–884
54. J. Crowe, B. Hayes-Gill, M. Sumner, C. Barratt, B. Palethorpe, C. Greenhalgh, O. Storz, A. Friday, J. Humble, C. Setchell, C. Randell, H.L. Muller, Modular sensor architecture for unobtrusive routine clinical diagnosis. Distributed Computing Systems Workshops, 2004. Proceedings of 24th International Conference , 2004, pp. 451–454
55. P. Kulkarni, Y. ztrk, Requirements and design spaces of mobile medical care. SIGMOBILE Mob. Comput. Commun. Rev. **11**(3), 12–30 (2007)
56. A. Jaimes, Sit straight (and tell me what I did today): A human posture alarm and activity summarization system. In Proceedings of the 2nd ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (Hilton, Singapore, November 11–11, 2005). CARPE '05. ACM, New York, NY, pp. 23–34
57. A. Milenkovic, C. Otto, E. Jovanov, Wireless sensor networks for personal health monitoring: Issues and an implementation. Computer Communications, Special issue: Wireless Sensor Networks: Performance, Reliability, Security, and Beyond (Elsevier, Amsterdam, 2006)
58. M. Sung, C. Marci, A. Pentland, Wearable feedback systems for rehabilitation. J. NeuroEng. Rehabil. **2**,17 (2005)
59. T. Hester, R. Hughes, D.M. Sherrill, B. Knorr, M. Akay, J. Stein, P. Bonato, Using wearable sensors to measure motor abilities following stroke. Wearable and implantable body sensor networks, 2006. BSN 2006. International Workshop, 3–5 April 2006
60. E. Jovanov, A. Milenkovic, C. Otto, P.C. de Groen, A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. J. NeuroEng. Rehabil. **2**, 6(2005). Available at: http://www.jneuroengrehab.com/content/2/1/6/
61. L. Meador (Panel Chairman), Leveraging basic research into MASINT capabilities (A five-year plan), Response to Conference Report to H.R. 4546, the National Defense Authorization Act for Fiscal 2003, March 2003
62. Research issues in data stream association rule mining N. Jiang, L. Gruenwald, ACM SIGMOD RECORD **35**, pp. 14–19 (2006)
63. S. Ivengar, S. Sastry, N. Balakrishnan, Foundations of data fusion for automation. IEEE Instrum. Meas. Mag. **6**(I4), 35–41 (2003)
64. L. O'Callaghan, N. Mishra, A. Meyerson, S. Guha, R. Motwani, Streaming data algorithms for high quality clustering. Proceedings of the 18th International Conference on Data Engineering, pp. 685–694, 2002
65. L. Weixian, L. Yilong, J. Fu, Data fusion of multiradar system by using genetic algorithm, IEEE Trans. Aerosp. Electron. Syst. **38**(2), 601–612 (2002)

66. M. Halatchev, L. Gruenwald, Estimating missing data in related sensor data streams. International Conference on Management of Data, Jan 2005, pp. 83–94
67. G.W. Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, M. Welsh, Deploying a wireless sensor network on an active volcano, In IEEE Internet Computing, Special issue on data-driven applications in sensor networks, March/April 2006
68. Z.K. Baker, V.K. Prasanna, Efficient Hardware Data Mining with the Apriori Algorithm on FPGAs, fccm, pp. 3–12, 13th Annual IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM'05), 2005